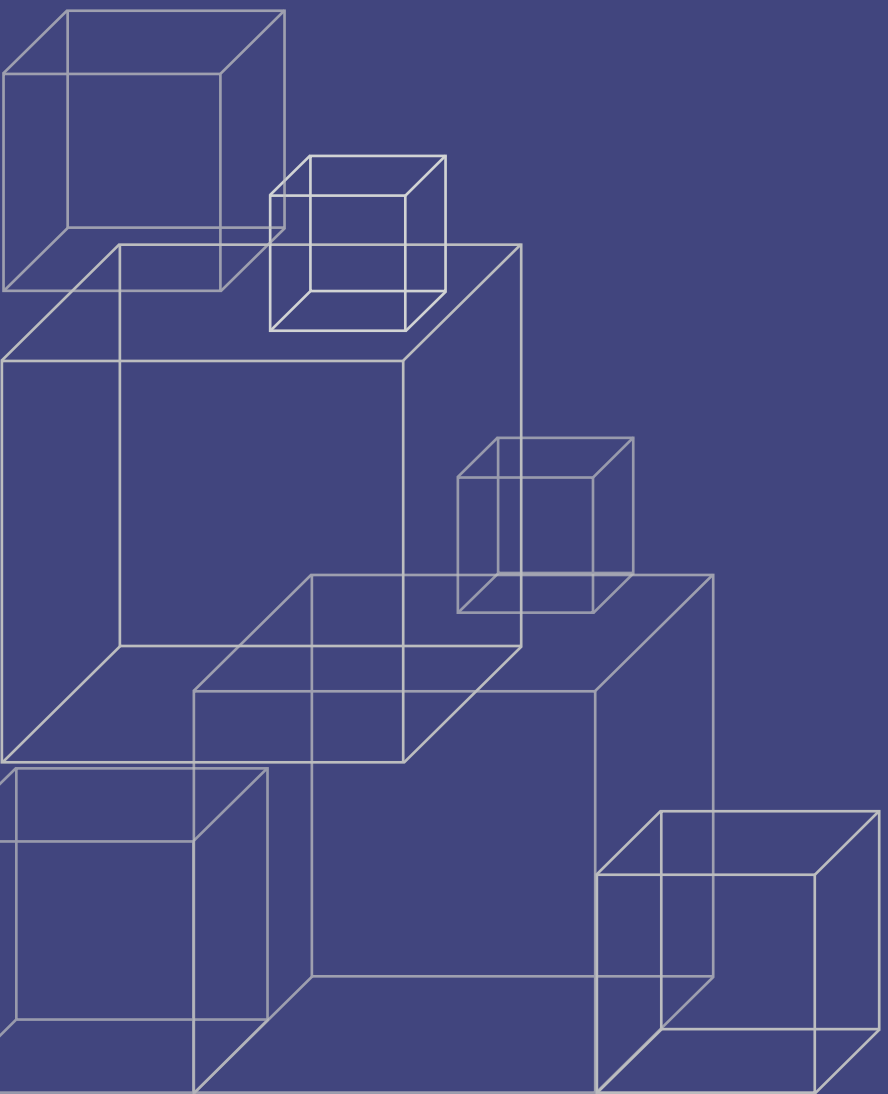
DATAQUALITY CAMPAIGN

Using Data To Improve Student Achievement

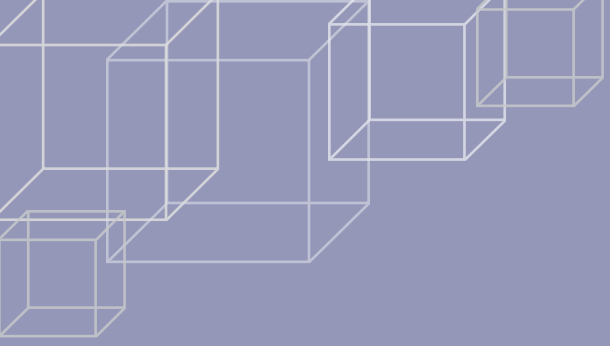
www.DataQualityCampaign.org

Creating Longitudinal Data Systems

Lessons Learned by Leading States



October 2006



This publication was produced by the Data Quality Campaign/National Center for Educational Accountability for submission under contract with the Bill & Melinda Gates Foundation. This report was written by Nancy J. Smith, Ph.D., Data Quality Campaign, and Jane Armstrong, Ph.D., director of Armstrong and Associates. Cherry Kugle, Ph.D., assisted in conducting the site visits and writing the state-specific reports.

Table of Contents

Improving Student Achievement

State and local leaders are under increasing pressure to improve student performance — to make certain that every student achieves academic proficiency and to close the achievement gap that has persisted for too many generations. The federal No Child Left Behind (NCLB) legislation further specifies that all students must be proficient in reading, mathematics and science by 2014. To encourage attainment of this goal, NCLB requires states to measure and report adequate yearly progress. To improve student achievement and meet NCLB reporting requirements, states will need to develop sophisticated, yet easy-to-use data systems that track the progress of individual students through their education and training lifetimes — from prekindergarten through postsecondary education and employment.

Longitudinal data systems account for individual students and make it possible to:

- follow students' academic progress as they move from grade to grade;
- determine the value-added and effectiveness of specific schools and programs;
- identify consistently higher-performing schools so that educators and the public can learn from best practices;
- evaluate the effect of teacher preparation and training programs on student achievement; and
- focus school systems on preparing a higher percentage of students for success in rigorous high school courses, college and challenging jobs.[1]

A Closer Look at Four Leading States

To help states accelerate their progress, the Data Quality Campaign (DQC) examined four diverse, leading states. The goal was to better understand how these states went about designing their data systems, what it cost to create them, what immediate and tangible results were achieved, and what "lessons learned" could be shared with other states following in their footsteps.[2] The four states are:

- **Florida,** which has been a pioneer of state data systems since it began its work in this area in 1986. Florida has a long and strong history of individual-level data collection systems in K–12 — and more recently, in higher education. The data system benefits from legislative support and funding.
- **Utah,** which developed its data system from scratch. Beginning with "stovepipe" data collections that were gathered into a clearinghouse, legislation in 2000 supported the creation of a data warehouse that integrates data from multiple systems and later added unique student identifiers.
- **Virginia,** which in 2000 began to fund development of a technology infrastructure in all 123 school divisions to support the state standards. By 2002, with legislative support, Virginia was one of the first states to use online testing and student identifiers for its student assessment program. Recent gubernatorial interest accelerated the completion of the longitudinal data system to answer education policy questions.

[1] Data Quality Campaign. *Creating a Longitudinal Data System: Using Data To Improve Student Achievement*, 2006.
[2] The four case studies (Florida, Utah, Virginia and Wisconsin) are available on the Web site of the Data Quality Campaign at www.DataQualityCampaign.org/state_specific/.

- **Wisconsin,** which faced a number of barriers in developing its system, including initial political opposition to a data system, concerns about compromising student privacy and a lack of funding. Fortunately, the NCLB reporting requirements and a strong state-level policy advocate provided impetus and weakened the opposition, while an external advisory committee helped design the system and pave the way for its development.

The lessons learned from these four case studies, which can be found in their entirety at www.DataQuality Campaign.org/state_specific/, are summarized in this issue brief. These guidelines are shared here to help state leaders who are now developing similar systems avoid the pitfalls and enhance the designs, implementation strategies and cost efficiencies of these four states.

Characteristics of Case Study States

	Florida	Utah	Virginia	Wisconsin
Number of districts and charter schools (as reported by state education agency staff)	67 districts 5 charters	40 districts 52 charters	132 districts 0 charters	425 districts 15 charters
Number of students	2.6 million	510,000	1.2 million	875,000
Inception of longitudinal data system	1986–87	2005–06	2005–06	2005–06
Legislative support and funding	Yes	Yes	Yes	No

Lessons Learned

Education leaders interviewed in the four case study states were generous in sharing their experiences. The following lessons learned seek to capture findings across the four states and organize them into a potential roadmap for creating longitudinal data systems.

Lay the Groundwork

Create public and political support for a longitudinal data system

Wisconsin leaders successfully used the data collection and reporting mandates from NCLB to build the necessary support to assign student identifiers — their first step in creating a longitudinal data system. Other states garnered legislative support by describing how a system could help improve student achievement and answer important policy questions — which led to funding and advocates for the system. Keeping key state leaders up to date on the development and use of the data system can maintain continued political and financial support for the system over election cycles.

Know your starting point

Identify what data systems are already in place in the state. Are student identifiers already available, or do they need to be created? Like Utah, learn from other states that have weathered the storms of creating student identifiers. Expect to spend a year in "discovery" (i.e., conducting a needs assessment, designing the project, defining the scope of the work and gathering stakeholder input) for a

Elements and Components of Longitudinal Data Systems

Framework for Interpreting the Lessons Learned from the Leading States

The DQC has identified *10 essential elements* that states must include to build a highly effective longitudinal data system:

1. A unique student identifier.
2. Student-level enrollment, demographic and program participation information.
3. The ability to match individual students' test scores from year to year to measure academic growth.
4. Information on untested students.
5. A teacher identifier system with the ability to match teachers to students.
6. Student-level transcript information, including information on courses completed and grades earned.
7. Student-level college readiness test scores.
8. Student-level graduation and dropout rates.
9. The ability to match student records between the P–12 and postsecondary systems.
10. A state data audit system assessing data quality, validity and reliability.

As of fall 2006, no state had a longitudinal data system that contained all 10 essential elements. To encourage and support the development of systems with all of these elements, the U.S. Department of Education awarded grants to 14 states in 2005.[3] The department expects to expand these grants to include additional states to support the building and use of longitudinal data systems. This issue brief provides policymakers with information to assist in this effort.

As outlined in *Creating a Longitudinal Data System: Using Data To Improve Student Achievement,*[4] a white paper by the DQC, a longitudinal data system needs to include the following components:

- **A technology infrastructure.** Schools, districts and state agencies have access to computers, servers, networks and the Internet to collect, transfer and use data.
- **A data architecture** that defines how data are coded, stored, managed and used. Data definitions are important. When everyone uses standard definitions, different systems can share information, staffing resources and process time are minimized, and data are provided to users when they need them. Privacy protection measures allow unique student identifiers to be used without revealing the data associated with a specific student when the data are shared with other organizations. Security protocols, like encryption, allow the secure transmission of data among systems.
- **A data warehouse** that stores, organizes and links student, school and district information — over time. Warehouses are designed to make it easy for users to "query" the database and produce standard or customized reports for different stakeholders. Researchers can use the data warehouse to answer questions such as the value-added of schools, identify which programs work for which students or identify which schools are closing the achievement gap — without violating student privacy.
- **Ongoing professional development** for those who are charged with collecting, storing, analyzing and using the data. Training ranges from how data are input locally to how teachers access and use the data for school and instructional improvement to how state education leaders use the system to make policy changes. Professional development continues as the system is refined and gains capacity for data-driven decisionmaking.

[3] The Institute of Education Sciences in the U.S. Department of Education provided three-year grants ranging from $1.6 million to $6 million to 14 states in 2005 to develop and enhance statewide longitudinal data systems. Go to http://nces.ed.gov/Programs/SLDS/stateinfo.asp for detailed abstracts of each state's grant.
[4] Data Quality Campaign, 2006.

large effort like a data warehouse, and know that there is no "shrink-wrapped" answer that will work for every state. Each state has to develop its own system based on its starting point, the desired functionality of the system and the available funding to support it.

For example, in 1986, Florida developed the Florida Information Resource Network to help districts transmit data to the state. A few years later, Florida implemented a student identifier system, an electronic transcript system and the ability to follow students after high school graduation. Combining the data from all of these systems into a data warehouse over the past few years has given Florida the ability to create more complete and integrated reports for the many users of education data throughout the state without duplicating efforts.

Involve state and district leaders as well as state education agency staff in the design of the system to meet their needs and increase its use

External groups should be involved to help identify what they need from the data system and how to use it. Data systems need enough flexibility to accommodate large districts that may already have a data system in place and also small districts that have no experience with and little resources/capacity for data systems. An internal advisory group should help define the data to be collected and the process for including changes and new features over time. A cross-departmental team within a state education agency should identify and coordinate reporting requirements to streamline data reporting for districts.

Utah was relentless in involving district leaders in the planning of its system, resulting in a strong data system that is used and supported by district staff. Florida sought advice from district leaders involved in the development and maintenance of its data systems. The state is now in the process of developing a new reporting tool called *Sunshine Connections* to ensure that local educators can access data easily and quickly, and it is continuing to involve the end user in the design of this resource.

Does your state have the data to answer these questions?

- Which schools produce the strongest academic growth for their students? (23 states report that they have the data to answer this question)*
- What achievement levels in middle school indicate that a student is on track to succeed in rigorous courses in high school? (5 states)
- How many students drop out — or are otherwise unaccounted for — after 8th grade? (40 states)
- What high school performance indicators (e.g., enrollment in rigorous courses or performance on state tests) are the best predictors of students' success in college or the workplace? (4 states)
- What percentage of high school graduates who go on to college take remedial courses? (14 states)

Design the System

Have a long-term implementation plan for the data system and the use of the data

Once a system is designed — based on stakeholder needs and factoring in existing components of the system — create a long-term plan that identifies when new components of the system will be available and how they will be shared by various users. Build the system in small, manageable and fundable phases; otherwise, the project will become too cumbersome for state and district personnel. Virginia has an effective 10-year plan and is carrying it out in two-year phases as funding is appropriated.

Anticipate future requirements of a longitudinal data system and design the initial system so new features can be added easily later

Visionaries in the case study states already are considering adding data system functions — such as "data marts"

*Although states report having the data elements to address this issue, the National Center for Educational Accountability/DQC survey did not capture whether states actually were using the data to this end.

for commonly requested reports, matching student records between P–12 and postsecondary institutions, linking teacher identifiers to student course outcomes, and linking student achievement data with coursework and strategies that can be used to improve instruction. Advisory councils are used to anticipate these requirements and ensure the design of the data system is flexible enough to accommodate them. With the U.S. Department of Education funding two states in 2006 to develop "growth models," requirements for these systems should be anticipated.[5]

Translate the design and stakeholder requirements into a well-specified request for proposals (RFP)

Florida began with a request for information from vendors to learn what was possible in designing their system. This allowed them to extract what they wanted in a system and create an invitation to negotiate with vendors. Utah developed an RFP and organized a panel to evaluate responses.

The selected vendor should have recommendations from other states and have deep staff expertise in creating educational data systems. Contracts should be written carefully to include dates; deliverables; and any special requirements, such as vendors working on site in states.[6]

Ensure that there is adequate funding to create the designed data system — and complete its implementation

States that have gone through the above steps should have a good estimate of the costs to create each part of a longitudinal data system. If accountability is built into the design and development process, then legislators — who hold the purse strings — can be assured they will get the system they paid for. Often it is valuable to communicate the cost of a data system as a percentage of education spending — and the value it will add to improving it.

Recognize and support district costs of data collection and implementation

Hold meetings with district leaders to avoid imposing too costly a set of changes. Invite district feedback on how to design and sequence the system for efficiency and effectiveness. Florida, for example, earmarks a percentage of state resources provided to districts for data and information systems.

Build and Maintain a Longitudinal Data System

Manage the expectations of policymakers, educators and the public throughout the design and implementation of the data system

First, be flexible in the face of changing leadership, technology, expectations and political environment. Also, expect the process to be difficult for both districts and the state for the first year or two as they work together to design the system, define the standards and get the kinks out of the system. Sharing the rollout plan and what data system functions can be expected — by when — will create support and use of the data as they become available. In the case study states, it took from 1.5 to three years to implement a student identifier system and a data warehouse. Although Virginia successfully created its system on an accelerated timeline, other states recommended taking more time to adequately plan and pilot the data system.

Assign dedicated staff and other experts to maintain as much continuity of leadership and staff as possible

Whether a state chooses to contract with a vendor or build the system itself, state education agency staff time will need to be dedicated to monitoring contracts, managing internal and external advisory committees, participating in and documenting the decision process, and consistently

[5] "Secretary Spellings Approves Tennessee and North Carolina Growth Model Pilots for 2005–2006." May 17, 2006. Available at www.ed.gov/news/pressreleases/2006/05/05172006a.html.
[6] Nancy J. Smith. *Lessons Learned: Writing Requests for Proposals for Statewide Student Data Systems in Education*. Education Commission of the States and InfoSynthesis, 2004.

communicating with all interested stakeholders to ensure the system will meet the needs of the end users. Over time, it is beneficial to build the capacity of state agency staff to design new features and maintain the data system — keeping expertise in house.

Encourage communication between the project team and the vendor, other departments within the state department of education, and the districts that will be both supplying and using the data

As with any project, effective and frequent communication is essential when building a longitudinal data system. This effort can be fraught with miscommunication if there is no shared understanding of common terms and expectations. The vendor and state staff may need to develop a common glossary of terms as a reference. Effective project managers facilitate ongoing communication among and input from all groups that are providing and using the data. This will ensure that the data system is easy to use and meets multiple stakeholders' needs.

Start with a focused effort that will create an "early win"

Longitudinal data systems are complex and require the coordinated design of many technology systems. Virginia started with grants to districts to create a technology infrastructure to implement the state academic standards. This was followed by online testing for the state assessment system. Each piece was celebrated as a "win," and district leaders saw the value of each of these components for improving student achievement. Keep the short-term scope small and have clear deliverables. Build the data system in phases and make sure the sequential components are working, rather than try to build all components at the same time.

Create an oversight committee to monitor and assess progress

State education agency staff in charge of building and maintaining the system must have the authority to make decisions or the capacity to get timely decisions from higher-level policymakers. Both internal and external oversight committees should be established to review progress and make higher-level decisions or recommendations both during the development phase and as updates and refinements are added over time.

Ensure the Accuracy and Integrity of the Data

The quality of the data is only as good as the data entered at the local level

Make sure that user training is thorough, the rules are clear, and sufficient assistance for district and school personnel is available. Experience in these four states and others shows that the quality of the data increases as they are used. At all levels of the system, people must take steps to ensure quality once they realize the data are being used for accountability, reporting and informing instruction.

Spend the time necessary to define and standardize the data elements and data collection processes before implementing the system

With district and vendor input, thoroughly document the data standards (e.g., definition of the element, acceptable values, how often it is to be collected, when it is to be collected, etc.). The more specific the standards for the data entry staff are, the more consistent and reliable the data in the state system will be.

Maintain security and confidentiality

Although building and using these longitudinal data systems are important for policy, management and instructional decisions that focus on individual student success, these needs must be balanced with appropriate protections for the privacy of student records. The Federal Educational Rights and Privacy Act (FERPA) imposes limits on the disclosure of student records by educational agencies and institutions. States must ensure they are collecting, sharing and using data in ways that comply with this federal law as well as their own state privacy laws, statutes and guidelines.

Some of the methods used to protect privacy include encrypting identifiable information (such as student identifiers), not reporting results for fewer than five or 10 students, and developing levels of password-protected access for data users. (For more information on how to build and manage state longitudinal data systems that protect individual privacy, please see *Maximizing the Power of Education Data while Ensuring Compliance with Federal Student Privacy Laws: A Guide for State Policymakers*, available at www.DataQualityCampaign.org.[7])

Promote the Use of the Data System and Be Flexible about Changes

Educate policymakers and educators about the data system and its potential uses to improve student achievement and support policy decisions

Marketing and public relations directed to state education agency staff, school districts, policymakers and the public are invaluable to making sure that stakeholders know about the advantages and usefulness of the system. Although *accountability* is the current driving force of data collection activities, Virginia noted that progressive districts see data collection as useful for *informing instruction* and targeting *intervention*.

States should build demand among the potential users of the longitudinal data as the system is built, and they should listen to what users need to know to ensure that the system can provide answers to their questions. Staff in Florida listen to the questions legislators are asking and provide applicable reports to them — whether or not official requests for information are made to the state education agency. Common users of longitudinal data include policymakers, district and school administrators, teachers, parents, students, and researchers.

Train educators, policymakers and researchers in the use of the system

Trainers should be senior-level staff with expert knowledge of the system. When district leaders in Utah began to see the benefits of the system, they developed a users group that facilitates communication with the state and shares knowledge and practices among districts.

Nurture the relationships between the people managing the student data and those managing the assessment data at the local level — so the data "fit together." It is helpful for districts to have their own data coordinators who can be liaisons to the state, as well as help local educators and administrators learn how to make regular use of the data. Training legislative aides would enable use of the system to answer policy questions and debates.

[7] Data Quality Campaign, 2006. Available at www.DataQualityCampaign.org/files/Publications-FERPA_A_Guide_for_State_Policymakers.PDF.

Encourage use of the system by creating a Web site with standard reports and query tools

Create user-friendly access to the most commonly requested reports, as well as the capacity to create custom reports. To generate demand for and use of data, the state needs to make the data easily, quickly and readily available to potential users. Web-based reporting and query tools can reach a large audience "on demand" — satisfying data needs quickly.

Create plans for ongoing maintenance and change

Information systems are never static; as federal and state laws change, accountability systems are improved and education programs are enhanced, the need to refine and alter the data system will arise. Stay close to the needs of the end users. Develop a process early on for managing and funding necessary changes and maintaining the system. Periodically review data elements to ensure they are the most useful ones to collect.

Learn from other states

Encourage system designers and users to attend national meetings such as the annual Management Information Systems conference and Summer Data Conference sponsored by the National Center for Education Statistics. This will help create awareness of what other leading states are doing. The DQC also provides forums for state policymakers and data managers to continue to learn from one another.

Estimating Costs of Longitudinal Data Systems

It is difficult to provide a single price tag for developing a state longitudinal data system because each system varies in scope and design. Furthermore, case study researchers found it difficult for states or districts to estimate the costs of collecting and using data because so much of the process is absorbed into the current infrastructure without financial support from the state. It is, however, possible to estimate costs for building specific components of the system.

With these caveats, best estimates from this study suggest a student identifier system can be built for between $1 million and $3 million. Some additional findings related to state costs:

- Each of the four case study states spent between $1 million and $3 million annually (while building the system over several years) for the various components of their systems — whether done internally, with contractors or through a vendor.
- Estimates of staff time range from the equivalent of six to 10 full-time staff members annually to oversee the state-level systems, meet state and federal reporting requirements, and provide support to districts.
- Costs for maintaining longitudinal data systems are also important to budget. Wisconsin estimated it spent $360,000 in information technology (IT) staff costs to maintain its system. Training costs were additional. The Utah Legislature appropriated $200,000 per year for maintenance.

Findings related to district costs:

- Florida reported that a portion of the legislative budget allocated to school districts must be used for data and information systems, and Virginia reported that districts may choose to use state appropriations designated to implement the state academic standards on data collection and submission systems. In the other states, districts do not receive state funds to offset data systems expenditures.
- District representatives indicated that most of the work associated with changing their data systems to accommodate the new state system was absorbed by having existing staff work overtime, delaying other projects and shifting responsibilities.
- Estimates of district staff time associated with maintenance and change to data systems range from the equivalent of eight to 10 full-time staff members in large districts and states to one or less than one full-time staff member in smaller districts and states. Most districts did not hire additional staff.

These costs do not take into account the real and potential savings that occur as a result of better data quality and the reduction of outdated and duplicate data collections. District and state representatives in every state indicated that although changing systems was difficult, the benefits outweighed the costs when considering the improved data quality and information available for research and decisionmaking.

Policymakers and state staff also should recognize that although building and implementing a longitudinal data system is costly and time consuming, it is not a one-time cost. The systems not only will need to be maintained in terms of hardware, software and annual training, but they also will need to be adapted over time to add and delete data elements as state and federal reporting requirements and accountability systems change. Instituting a detailed process for the annual review of data elements, data collection procedures, training methods and infrastructure upgrades should be a part of the state's long-term vision of the data system.

Specific Costs by State

- **Florida:** Most development costs for Florida's system were borne in prior years. However, a *quid pro quo* was negotiated with a vendor to develop the data warehouse. Currently, six full-time staff in the state education agency provide programming support to local school districts.
- **Utah:** For 2005 and 2006, it is estimated that state-level IT costs totaled $800,000 per year (to create the student identifier, data warehouse, clearinghouse, NCLB reporting, and test scanning and scoring). Each year, approximately 10 full-time state education agency staff supported the state-level data system.
- **Virginia:** In 2000, $3.6 million was appropriated to provide funds to districts to build a technology infrastructure — including Internet-ready local area networks and high-speed, high-bandwidth capability in all schools. The statewide information system and data warehouse cost approximately $3 million to maintain — plus staff time.
- **Wisconsin:** Approximately $650,000 was contracted to a vendor to develop the student identifier system, and another $650,000 was used for student-level enrollment data collection. This two-year effort was supplemented by about $1.3 million in state education agency IT staff time.

Key Questions To Consider

All 50 states are at very different places in the effort to build, manage and use state longitudinal data systems. The results of the annual DQC/National Center for Educational Accountability survey of state data collection issues related to longitudinal analysis, available at www.DataQualityCampaign.org, highlight the fact that most states are focusing on building, growing and improving these systems. To inform and promote continuous strengthening of these state systems, the DQC offers the following policy, design, implementation and support questions for state and local leaders to consider:

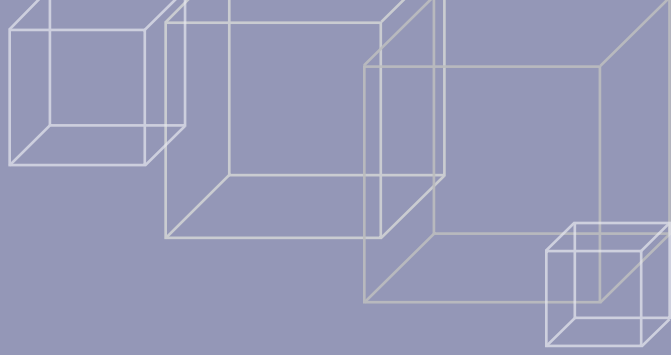
- What data system do you want? Remember the 10 essential elements proposed by the DQC to create a robust data system.
- Given these elements, what policy and school improvement questions do you want your system to answer? How can you involve potential users in the design and use of the data system?
- What are the design specifications? How can you answer the data needs of state and local policymakers — as well as the public?
- What technology and data infrastructures do you already have in place that you can build on?
- What expert assistance do you need? How can you ensure these vendors are expert, are cost effective and will deliver what they promise? How do you hold them accountable?
- How should local educators be involved and trained in these new systems — since they both originate and use the data?
- How can states both create and be responsive to feedback about user needs?
- Given that the ultimate goal of these systems is to improve student achievement, what is the state role, what is the district role and what supports do they need to use these data to improve student achievement?
- What are new uses for the data system? How do you identify these new uses and ensure the system is flexible enough to include them easily?
- How does the state longitudinal data system work with/complement existing district data systems?
- How do you guarantee adequate ongoing resources to maintain and enhance the system?

The responsibility to improve student achievement is universal. State, district and school leaders, as well as school boards, parents and other stakeholders, have a role to play in improving student achievement. It is impossible to know where to focus improvement efforts without a data system that identifies student needs as well as the policies, investments and practices that have been proven to improve achievement. Longitudinal data systems will provide this information. The rest is up to us.

To Find Out More

Visit www.DataQualityCampaign.org to find out more about:

- the 10 essential elements;
- the case studies for Florida, Utah, Virginia and Wisconsin; and
- a guide for policymakers on maximizing the power of longitudinal data while protecting student privacy.



DATAQUALITY CAMPAIGN

Using Data To Improve Student Achievement

The Data Quality Campaign is a national, collaborative effort to encourage and support state policymakers to improve the collection, availability and use of high-quality education data and to implement state longitudinal data systems to improve student achievement. The campaign aims to provide tools and resources that will assist state development of quality longitudinal data systems, while also providing a national forum for reducing duplication of effort and promoting greater coordination and consensus among the organizations focusing on improving data quality, access and use.

The campaign is managed by the National Center for Educational Accountability and supported by the Bill & Melinda Gates Foundation.

www.DataQualityCampaign.org

Managing partners of the Data Quality Campaign include:

- Achieve, Inc.
- Alliance for Excellent Education
- Council of Chief State School Officers
- Education Commission of the States
- The Education Trust
- National Association of State Boards of Education
- National Association of System Heads
- National Center for Educational Accountability
- National Center for Higher Education Management Systems
- National Governors Association Center for Best Practices
- Schools Interoperability Framework Association
- Standard & Poor's School Evaluation Services
- State Educational Technology Directors Association
- State Higher Education Executive Officers

Endorsing partners of the Data Quality Campaign include:

- ACT
- Alliance for Quality Teaching
- American Association of Colleges for Teacher Education
- American Association of State Colleges and Universities
- American Board for Certification of Teaching Excellence
- APQC
- Center for Teacher Quality
- College Summit, Inc.
- Consortium for School Networking
- Educational Policy Institute
- GreatSchools
- Jobs for the Future
- League of Education Voters Foundation
- National Alliance for Public Charter Schools
- National Association of Secondary School Principals
- National Education Knowledge Industry Association
- Postsecondary Electronic Standards Council
- Roads to Success
- Southern Regional Education Board